Kamala Harris' Popular Vote Support in the 2024 US Election*

Justin Klip Dhruv Gupta

October 22, 2024

First sentence. Second sentence. Third sentence. Fourth sentence.

1 Introduction

Why it matters paragraph: Owing to the USA's economic prowess, the American election every 4 years is one of, if not the most, important election in the modern world. Due to this, political scientists and pollsters work tirelessly in order to provide an accurate forecasting of the election. That being said, in 2016 countless pollsters were unable to correctly predict it's eventual winner. This raised questions about the legitimacy of polls and made me pollsters reevaluate what metrics they use to predict elections (Bokat-Lindell 2020).

Overview paragraph: In this paper, we generate a model using a poll of polls approach outlined in (Blumenthal 2014) order to predict popular support for Kamala Harris in the 2024 US election.

Estimand paragraph: The estimand of interest in this paper is simply the popular support in percentage of Kamala Harris for the 2024 paper. With this popular support model built, other researchers can build and try to extend this model to look at electoral college forecasts.

Results paragraph:

Telegraphing paragraph: The remainder of this paper is structured as follows. Section 2 provides an overview of how the data was obtained and cleaned. It also provides a detailed section on measurement and our outcome variables of interest. Section 3 describes the setup of the model and the rationale behind it. Section 5 interprets the results and their validity.

*Code and data are available at: https://github.com/justinklip/usa-election-forecast-2024.

2 Data

2.1 Overview

We use the statistical programming language R (R Core Team 2023) to run our analyses. Our data comes from 538's general presidential poll data for the 2024 election cycle. It contains all large presidential controls that meet certain criteria (this is discussed in **?@sec-measurement**) (FiveThirtyEight 2024). Following Alexander (2023), we consider use R packages to clean the data such as:

Overview text

2.2 Measurement

Since this paper makes use of multiple polls and polling types, there is a lot of variance in how the data in this data set may measure their data. They do, however, generally fall under the same principles of polling when trying to obtain this data.

Firstly, a pollster (such as YouGov, Siena/NYT and others) will conduct the poll in the real world. We document how one pollster (YouGov) conducts their poll in the specific in the Section B section. When this poll is complete, FiveThirtyEight collects the polling data from these pollsters. They collect any publicly available data that meets the following criteria: the poll has the name of the pollster, survey dates, sample sizes, and details of population sample. The data must also answer questions about their medium of survey, source of files, weighting, and funding (FiveThirtyEight 2020).

2.3 Outcome variables

2.4 Predictor variables

3 Model

3.1 Model set-up

We run the model in R (R Core Team 2023) using the `rstanarm` package of Goodrich et al. (2022). We use the default priors from `rstanarm`.

3.1.1 Model justification

4 Results

5 Discussion

5.1 First discussion point

5.2 Second discussion point

5.3 Third discussion point

5.4 Weaknesses and next steps

Appendix

A Additional data details

This paper does not have an accompanying download for the data script, but the data was from FiveThirtyEight (2024). It was downloaded by navigating the website in the citation, filtering to "President: General Election", "National" and "2024" for poll type, state, and cycle respectively. We then pressed download and saved the csv to our raw data folder.

B Pollster Methodology and Evaluation

C Model details

C.1 Posterior predictive check

C.2 Diagnostics

References

Alexander, Rohan. 2023. *Telling Stories with Data*. Chapman; Hall/CRC. https://tellingstorieswithdata.com/.

Blumenthal, Mark. 2014. "Polls, Forecasts, and Aggregators." *PS: Political Science & Politics* 47 (2): 297–300. https://doi.org/10.1017/S1049096514000386.

Bokat-Lindell, Spencer. 2020. "Opinion | How to Think about Election Forecasts." *The New York Times*. https://www.nytimes.com/2020/11/01/opinion/election-forecasts-modeling-flaws.html.

FiveThirtyEight. 2020. "Polls, Policy, and FAQs." https://fivethirtyeight.com/features/polls-policy-and-faqs/.

———. 2024. "2024 National President General Election Polls." https://projects.fivethirtyeight.com/polls/president-general/2024/national/.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "rstanarm: Bayesian applied regression modeling via Stan." https://mc-stan.org/rstanarm/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.